CHAPTER 12.

Accuracy Assessment of Insect and Disease Survey and eDaRT for Monitoring Forest Health

Michèle R. Slaton
Kayanna Warren
Alexander Koltunov
Sheri Smith

INTRODUCTION

The historic drought and bark beetle outbreak in California that peaked 2014–2016 resulted in high levels of tree mortality at a landscape scale. Millions of trees were reportedly affected, primarily ponderosa pine (*Pinus ponderosa*), but also Jeffrey pine (*P. jeffreyi*), sugar pine (*P. lambertiana*), white fir (*Abies concolor*), and red fir (*A. magnifica*) (USDA Forest Service 2020b). U.S. Department of Agriculture Forest Service, State and Private Forestry, Forest Health Protection (FHP) delivers forest health information, including Aerial Detection Survey (ADS) data in support of the national Insect and Disease Survey (IDS). Aerial Detection Survey polygons are sketch-mapped from aircraft and attributed with disturbance agent and host and with estimated trees per acre (TPA) with mortality. Focus is typically on the upper forest stratum, defined for our study in the Sierra Nevada as conifers >11 inches diameter at breast height (d.b.h.).

As tree mortality progressed in 2014, high-quality, near real-time data were needed to facilitate timely management response. Toward this end, FHP undertook an evaluation of potential improvements to ADS, including integration of alternate technologies. Recently, satellite-based remote sensing methods and processing systems have emerged to detect stand- and landscape-level changes, including the Ecosystem Disturbance and Recovery Tracker (eDaRT; Koltunov and others 2020). eDaRT products provide information on vegetation disturbance events derived from Landsat satellite imagery at the 98-foot (30-m) scale. The onset of these events is attributed to a 2-week time period, and each event has a relative intensity that corresponds to canopy cover loss. The eDaRT system is both flexible and efficient for operational implementation needs, has demonstrated high accuracy (Koltunov and others 2020), and is in use for several applications in California.

Although both eDaRT and ADS may be used to represent forest mortality information, each one reports different metrics at different spatial and temporal scales. The overall goal of our Evaluation Monitoring project (EM-18-WC) funded by the Forest Service, Forest Health Monitoring program was to provide a rigorous assessment of ADS and eDaRT products and outline recommendations for complementary use by land managers. To meet this goal, our specific objectives were to (1) develop and execute a ground-based sampling strategy to compare ADS and eDaRT to reference validation data, and (2) make recommendations for appropriate use of data in varied applications.

METHODS

Our study area (104,795 acres) was located on the Sierra National Forest (SNF) in ponderosa pine and mixed conifer forests at elevations between 2,800 and 7,600 feet (fig. 12.1). Our sample consisted of 57 total ADS polygons (Damage Type 2 for mortality) and eDaRT (version 2.9) outputs for the entire study area

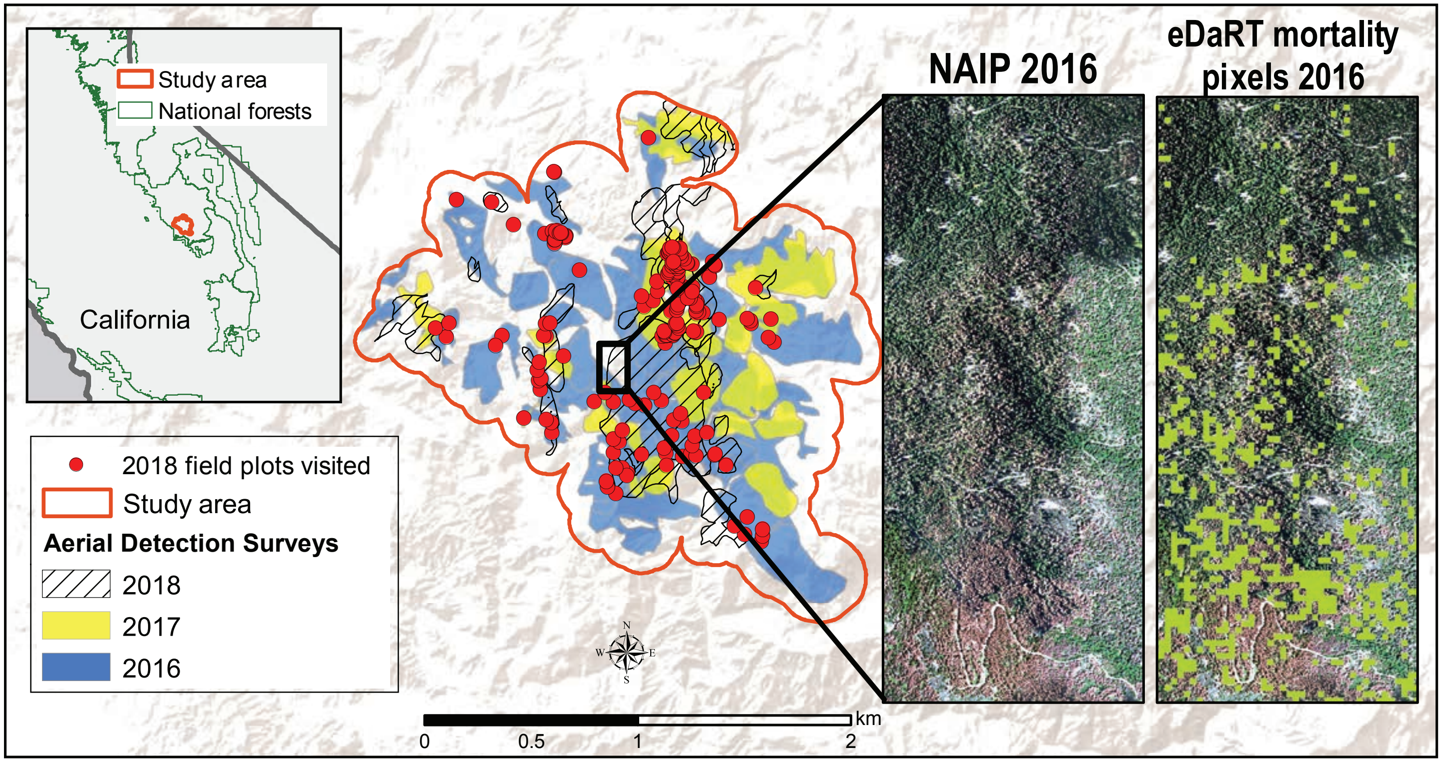

Figure 12.1—Map of study area on the Sierra National Forest, CA, including field plots and Aerial Detection Survey (ADS) polygons (years 2016 and 2018 only used in analysis).

for 2016 and 2018, years for which several ADS mortality polygons were mapped in the study area. Our analysis focused on ADS locality and relative severity data, which could be compared to eDaRT. All field sampling occurred in 2018, and our design targeted a primary sampling area based on *a priori* information about tree mortality occurrence (fig. 12.1). Within this area, we sampled 161 Landsat pixels (98 x 98 feet) within ADS polygons, and 21 pixels outside polygons, all randomly selected within forest types, excluding known fires or vegetation management projects, and prioritized by accessibility. Field technicians recorded tree needle color and decay class, as based on fine twig, bole, and bark characteristics to estimate year of death as presumptively recorded by aerial survey. Decay classes followed Forest Service procedures (USDA Forest Service 2020a), and needle color characteristics were based upon

Miller and Keen (1960) and further refined for the Sierra Nevada montane belt based on input from Forest Service pathologists and entomologists (key available upon request).

At the time of this project, the primary metric for magnitude of change for eDaRT events was a relative proxy for canopy cover loss, based on the temporally integrated standardized residuals of modeling vegetation indices (e.g., normalized difference vegetation index [NDVI]) before and after tree mortality events, a metric termed "confidence" (Koltunov and others 2020). We used field and imagery inspections to identify a threshold for spectral change, above which pixels were determined to be significantly affected by tree mortality, a class hereafter referred to as mortality pixels (eDaRT confidence >50). This relatively simple calibration would not necessarily apply elsewhere in the State, but at the time of this project, the threshold provided a simple and consistent method to deliver an estimate, while more robust estimators of event severity are under development in eDaRT. It is important to note that eDaRT products do not directly describe the number of trees, live or dead, but the degree to which each pixel is affected. Therefore, tree density products were created using F3, an interpolation algorithm developed at the Forest Service Pacific Southwest Region Remote Sensing Lab that uses Forest Inventory and Analysis (FIA) data and Landsat imagery to map forest stand metrics (Huang and others 2018). By overlaying F3 density of trees (prior to mortality) with eDaRT mortality products, we derived estimates of dead tree numbers from this combined eDaRT-F3 product. Resulting values are single estimates per pixel, with error rates for F3 and eDaRT as reported by Koltunov and others (2020) and Huang and others (2018), respectively.

Total dead tree numbers as derived from plots, ADS, and eDaRT were computed at the scale of individual ADS polygons and across the entire study area. For the plots, we multiplied mean density of dead trees by respective area acreage to determine the plot-based total for dead trees. For ADS data, we multiplied the dead TPA for each polygon by its acreage to derive total dead tree numbers. eDaRT mortality pixels were combined with F3 tree density data to compute a combined eDaRT-F3 result for the total quantity of dead trees. Error matrices comparing ADS to plot-based TPA estimates were created by computing means of dead tree densities for field plots within each polygon and binning values into TPA categories in increments of 10. User's, producer's, and overall accuracy were calculated for ADS following Pontius and Millones (2011). Logistical difficulty in precisely colocating ground plots with Landsat pixels led us to rely upon the polygon-scale comparisons of eDaRT to plot-based and ADS data, and also upon the high-resolution imagery-based assessment as reported by Koltunov and others (2020) for pixel-based error rates.

RESULTS

For 31 out of 36 ADS polygons mapped in 2016 (one of 37 polygons was discarded due to an apparently mislabeled TPA value of zero), the plot-derived densities of dead trees in the polygons (tallied in 2018; red phase; decay classes 2, 3, and 4; i.e., estimated 2–4 years since mortality) were lower than ADS data (fig. 12.2). For year 2018 ADS polygons, plot-based TPA estimates (tallied in 2018, red or yellow phase, decay classes 1 and 2) were also lower than ADS assignments in 17 out of 19 polygons. User's and producer's accuracies, representing omission and commission errors for each TPA class, ranged from 0–50 percent, and overall accuracies for 2016 and 2018 report years were 3 and 10 percent, respectively (fig. 12.2). Because of the difficulty in assigning year of tree death in the field, we allowed for a 1–2-year systematic misassignment by field technicians (i.e., up to three decay classes to represent the 2016 ADS report year, and two decay classes for 2018) but still found that ADS routinely over-estimated TPA. When compiling all TPA bins and examining accuracy only in terms of presence/absence of mortality in polygons, we found overall 40-percent accuracy for 2018 and 75-percent accuracy for 2016.

Across the study area, we found 48,414 acres in year 2016 ADS mortality polygons, in contrast to 7,048 acres as estimated by eDaRT. At the extent of the entire SNF, ADS mapped 181,725 acres with mortality in 2016 compared to 51,298 acres mapped by eDaRT, demonstrating a consistent pattern of greater area estimates from ADS compared to eDaRT. Aerial Detection

(A) 2018

ADS TPA class	Plot TPA class: 0	1–10	11–20	>20	Grand Total	User's accuracy (%)
0	0	1[a]	0	0	1	0
1–10	6	2	0	0	8	25
11–20	2	3	0	0	5	0
>20	3	3	0	0	6	0
Grand total	11	9	0	0	20	
Producer's accuracy (%)	0	22	0	0		Overall accuracy = 10%

[a] Indicates one apparently mislabeled ADS polygon with TPA value of zero.

(B) 2016

ADS TPA class	Plot TPA class: 0	1–10	11–20	>20	Grand Total	User's accuracy (%)
0	1	1[a]	0	0	2	50
1–10	1	0	0	0	1	0
11–20	4	2	0	0	6	0
>20	20	8	0	0	28	0
Grand total	26	11	0	0	37	
Producer's accuracy (%)	4	0	0	0		Overall accuracy = 3%

[a] Indicates one apparently mislabeled ADS polygon with TPA value of zero.

Figure 12.2—Error matrices representing disagreement between Aerial Detection Survey (ADS) and plot-based dead trees per acre (TPA) in polygons. Cell values are numbers of ADS polygons. Plot-based trees were >11 inches d.b.h. and included decay classes 1 and 2 for year 2018 (A) and decay classes 2, 3, and 4 for year 2016 (B). Raw values only are displayed; an area-based matrix with corrections following Pontius and Millones (2011) produced almost identical values.

Survey estimated a total of 1,954,086 dead trees in 2016 across our study area. This is in contrast to the plot-based estimate of 154,453 total dead trees (measured in 2018 as red phase decay class 2, >11 inches d.b.h.). A temporal allowance of +/- 2 years for potential misassignment of time of death by field technicians would be required to bring ADS estimates within the plot-based range of values for both flight years 2016 and 2018 (data not shown). Given the criteria we used to assess timing of death, and the fact that the criteria were largely developed for disease agents and hosts within our study area, it is unlikely that misassignment in timing of death explains the large departures between datasets. The combined eDaRT-F3 data indicated that the study area contained 488,945 dead trees in 2016. In both years 2016 and 2018, eDaRT-F3 estimates fell within the range of plot-based data given +/- 1-year error allowed for field assignment in time of death (fig. 12.3).

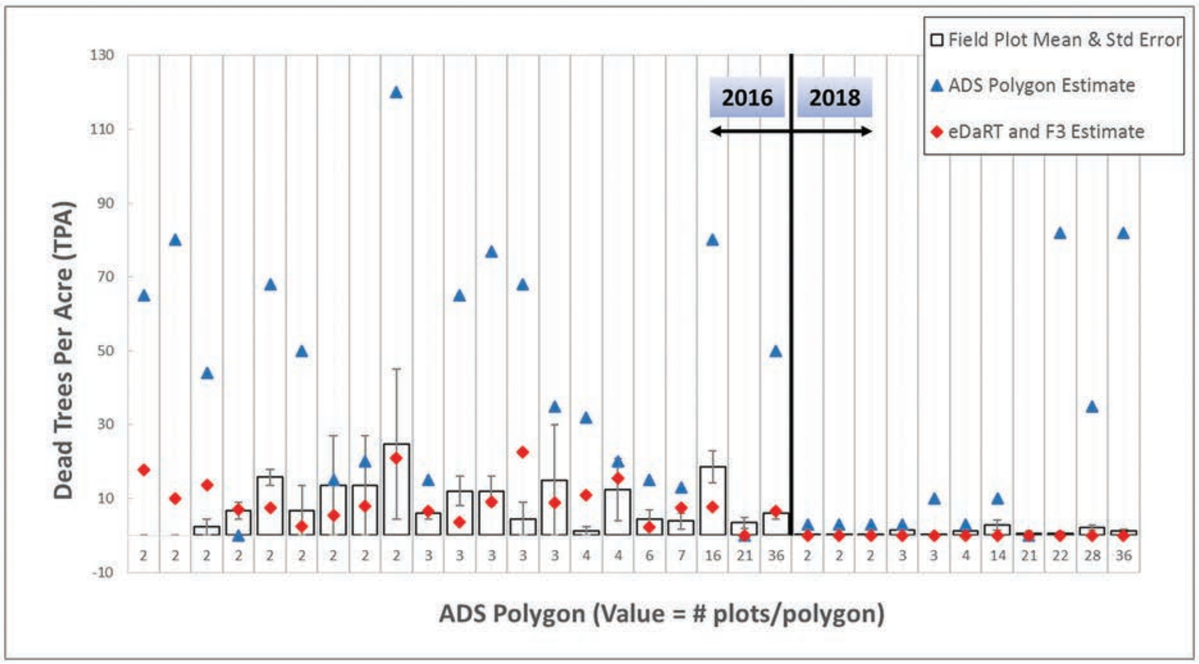

Figure 12.3—Dead trees per acre (TPA) estimates from ADS polygons, eDaRT-F3, and plot-measured means with standard error (SE) bars, with plot sample size per polygon for 2 years.

DISCUSSION

We generally found fewer dead trees within plots than corresponding ADS polygons indicated, both for 2016 and 2018 report years. Even given a temporal margin of error for field personnel assigning time of tree death of up to 2 years, we found that overall, plot-based estimates matched polygon estimates in TPA bins in only 3–10 percent of cases. Notably, a previous assessment of ADS across multiple States in the United States during less severe mortality events found a general pattern of underestimation of dead TPA as compared to ground- and imagery-based measurements (Coleman and others 2018). In our study area, extremely severe mortality occurred over numerous years, and in numerous species with inconsistent crown fade patterns. As a result, sketch-mapping required representation of complex patterning. Aerial Detection Survey practitioners report that in large events such as this, observations of dead tree patches are “lassoed” into larger and potentially more diffuse polygons to facilitate consistent tracking during flight; practitioners typically refer to polygons as areas “with” (i.e., that contain) mortality rather than “of” (i.e., totally affected by) mortality to make this distinction. Our findings were consistent with that trend, in which polygons with high TPA labels contained areas without dead trees. At the scale of the study area and of the SNF administrative unit, acreages of eDaRT mortality were generally lower than those reported by ADS. This may be partially due to the “lasso” effect of mapping extensive mortality that essentially lumps together several areas of acres *with* mortality, rather than splitting more tightly defined acres *of* mortality. Accordingly, FHP specifies in ADS metadata that products should be utilized at spatial scales of 1:100,000 or smaller and that overlays with other spatial products should be approached cautiously (USDA Forest Service 2020b).

Some discrepancy may also be explained by missed detections of low-magnitude mortality in eDaRT. However, levels as low as 5-percent cover loss have up to a 56-percent chance of detection, making missed detections relatively uncommon, especially in denser forests (Koltunov and others 2020). Another source of disagreement is likely rooted in each dataset’s method to identify timing of tree death. eDaRT timing is identified by the date of the first Landsat image (every 8–16 days) in which an anomaly is detected, and which may include invisible changes, such as canopy drying. Aerial Detection Survey mappers use visual cues based on canopy color to determine the year of mortality onset. However, because the severity of drought advanced the rate of canopy change, and because mortality expanded and intensified over numerous years, identifying precise timing of mortality onset from aircraft during this extreme event was challenging. There is also error associated with back-dating mortality on the ground given different fade patterns by species and in different microclimates; however, our 2–3-year temporal allowance for error in plot-based assignments reduces the likelihood that this factor explains the discrepancy. Lastly, eDaRT identifies disturbance dates based on the first image in which an anomaly is expressed.

However, if clouds or other factors compromise image quality, detection onset may be delayed, in some cases even until the following year.

The difference in total numbers of dead trees between ADS and the eDaRT-F3 across the study area in part reflects the differences in the areal extent (acreage) of mortality between the datasets. However, our findings show that the disagreement also stemmed from the ADS TPA estimate itself, which was greater than both the plot-based and eDaRT-F3-based values. Importantly, at the time of our analysis, ADS polygons were not labeled with an estimate of the proportion of the polygon (percentage of area) affected by mortality; this method was implemented in 2019. Additional validation will be required for this new metric, but it may be expected to increase the accuracy of derived metrics for numbers of dead trees from ADS because it potentially corrects for overestimation caused by the "lasso" effect.

Although deriving total numbers of dead trees from either ADS or eDaRT is possible, both methods should be approached with caution. Because tree numbers are not easily counted from aircraft, nor are they strongly correlated to spectral indicators in imagery, alternative mortality metrics, such as basal area, volume, biomass, or percentage of area affected may provide more reliable indicators for the magnitude of change. The former two metrics are currently available using eDaRT and F3, and the latter is a standard attribute for ADS products since 2019.

CONCLUSIONS

Practically speaking, forest health information must be accurate enough to guide allocation of funds and resources at a subregional scale and to plan and implement project prescriptions at the stand scale. Table 12.1 provides a simple decision-making matrix to assist users in data selection for determining tree mortality levels for land management applications and analyses. This matrix reinforces the importance of recent direction in the IDS program to enhance integration between aerial detection, ground-based assessments, and a range of remote sensing techniques, such as eDaRT and F3. Recommendations for future development and assessment of ADS and eDaRT methods are summarized into a few main points:

- The new ADS label for the percentage of a polygon affected is expected to improve the accuracy and utility of ADS dead tree products, although it will require both validation and quality control to reduce bias between observers. Similarly, the confidence-based threshold used to estimate severity of eDaRT events will be improved upon through the upcoming release of a Mortality Magnitude Index to directly estimate canopy cover loss, and F3 products will be improved by the recent implementation of intensification of remeasurement from a 10-year to a 5-year cycle.
- Demand for reference validation datasets remains high. "Virtual plots" in which analysts measure canopy cover change within Landsat pixels using high-resolution

Table 12.1—A decision-making matrix to assist users in data selection for determining tree mortality levels for land management applications and analyses

Information needed	Suitability of data currently available	Recommended developments
Agent and host tree species affecting area	ADS attributes for host type affected and agent are generally highly accurate (refer to Coleman and others [2018] and citations therein for assessment).	An eDaRT mortality-type classifier is currently under development, and ADS and ground plots may assist in training.
Acreage of tree mortality events	ADS is available but subject to error as reported in this chapter. ADS reports acres with mortality rather than acres of mortality.	Validation of ADS percentage of area affected metric
	eDaRT is appropriate from stand scale to State scale but relies on proxy for magnitude of events.	eDaRT Mortality Magnitude Index to quantify canopy cover loss is planned to replace proxy method (data release targeted for 2021).
Number of trees affected at district or broader scale	ADS is available but subject to error as reported in this chapter.	Validation of ADS percentage of area affected metric Investigation of dependency of accuracy on polygon size and proximity to polygon perimeter
	eDaRT-F3 was found to be highly accurate in our study area.	Additional eDaRT-F3 validation is recommended elsewhere prior to use.
Basal area or volume affected by mortality event	ADS may be used with standard tree allometry to derive these metrics, but no validation is available, and is subject to potential overestimation as found for number of trees.	Metrics are best suited to imagery and LiDAR-based approaches and are not likely to be developed in ADS.
	Combine eDaRT with F3; validation is not available, but basal area and volume are generally more readily modeled than number of trees.	Combined eDaRT and F3 products require additional validation for these metrics.
Settings and situations most appropriate for current use	ADS highlights broad-scale and annual trends, and detects small-scale and highly dispersed or light-severity events where agent and host identification are important. eDaRT-F3 is suitable for stand-level to regional analysis and currently best detects tree mortality in denser stands. eDaRT is also suitable for applications where high temporal resolution for disturbance onset is required.	The current study mainly compared ADS and eDaRT in the context of a severe and discrete mortality event. Continued exploration of product integration to exploit the strengths of each will be highly valuable.

ADS = Aerial Detection Survey; eDaRT = Ecosystem Disturbance and Recovery Tracker; F3 = an interpolation algorithm that uses Forest Inventory and Analysis data and Landsat imagery to map forest stand metrics.

imagery such as National Agriculture Imagery Program (NAIP) or from Google Earth provide an efficient means to create reference data for eDaRT (Koltunov and others 2020). Statewide LiDAR acquisitions, especially in conjunction with high-resolution imagery to assess crown fade and moisture, will provide for continued broad-scale assessment of both ADS and eDaRT. Implementation of these remote sensing-based methods will allow for improved targeting and efficiency of smaller scale ground-based validation efforts.

- Leveraging the unique strengths of diverse monitoring methods is becoming increasingly important as broad-scale tree mortality events unfold. Directing ADS flights toward core areas for targeted host and agent assessment will not only minimize risk for flight personnel but will complement remote sensing techniques by helping to train broad-scale models and to direct on-the-ground assessments.

For more information (including full EM-18-WC report), contact: Michèle Slaton, michele.slaton@usda.gov.

ACKNOWLEDGMENTS

We are grateful to Carol Clark for assistance with this analysis and to Adrian Das, Derek Young, Ryan Hanavan, and Jeff Moore for their comments that improved upon earlier versions of this manuscript.

LITERATURE CITED

Coleman, T.W.; Graves, A.D.; Heath, Z. [and others]. 2018. Accuracy of aerial detection surveys for mapping insect and disease disturbances in the United States. Forest Ecology and Management. 430: 321–336. https://doi.org/10.1016/j.foreco.2018.08.020.

Huang, S.L.; Ramirez, C.; McElhaney, M.; Evans, K. 2018. F-3: simulating spatiotemporal forest change from field inventory, remote sensing, growth modeling, and management actions. Forest Ecology and Management. 415: 26–37. https://doi.org/10.1016/j.foreco.2018.02.026.

Koltunov, A.; Ramirez, C.M.; Ustin, S.L. [and others]. 2020. eDaRT: The Ecosystem Disturbance and Recovery Tracker system for monitoring landscape disturbances and their cumulative effects. Remote Sensing of Environment. 238: 111482. https://doi.org/10.1016/j.rse.2019.111482.

Miller, J.M.; Keen, F.P. 1960. Biology and control of the western pine beetle. Misc. Publ. 800. Washington, DC: U.S. Department of Agriculture Forest Service. 318 p.

Pontius, R.G.; Millones, M. 2011. Death to Kappa: birth of quantity disagreement and allocation disagreement for accuracy assessment. International Journal of Remote Sensing. 32: 4407–4429. https://doi.org/10.1080/01431161.2011.552923.

U.S. Department of Agriculture (USDA) Forest Service. 2020a. Forest Inventory and Analysis national core field guide volume 1: field data collection procedures for Phase 2 plots, version 9.0. Washington, DC: U.S. Department of Agriculture Forest Service. 449 p.

U.S. Department of Agriculture (USDA) Forest Service, Pacific Southwest Region, Forest Health Protection. 2020b. Aerial Detection Monitoring. https://www.fs.usda.gov/detail/r5/forest-grasslandhealth/?cid=fsbdev3_046696. [Date accessed: June 30, 2020].